



Электронные библиотеки: проблемы и приоритетные направления развития

О. В. Сюттюренко

*Всероссийский институт научной и технической информации Российской академии наук,
г. Москва, Россия*

Аннотация: С системных позиций рассмотрены основные факторы-детерминанты, негативно влияющие на развитие электронных библиотек (ЭБ) и использование электронной научно-технической информации в целом в глобальной сети. Показаны подходы к созданию механизма смысловой навигации и поиска знаний в информационных сетях. Обозначены наиболее приоритетные и актуальные направления и задачи в области повышения эффективности разработки и использования ЭБ. Отмечается, что организация работ по созданию ЭБ – это междисциплинарная задача, требующая консолидации сил специалистов в целевых предметных областях, в области информационных технологий, управления и права.

Ключевые слова: электронные библиотеки; информационная среда; риски; информационные ресурсы; сетевая навигация и поиск; аналитическая постобработка; вебметрия; информационная инфраструктура

Для цитирования: Сюттюренко О. В. Электронные библиотеки: проблемы и приоритетные направления развития // Научное издание международного уровня – 2019: стратегия и тактика управления и развития: материалы 8-й Международ. науч.-практ. конф., Москва, 23–26 апреля 2019 г. Екатеринбург: Изд-во Урал. ун-та, 2019. С. 101–107. DOI: 10.24069/konf-23-26-04-2019.15.

Electronic libraries: problems and priority directions of development

O. V. Syntyurenko

*All-Russian Institute of Scientific and Technical Information of the Russian Academy of Sciences,
Moscow, Russia*

Abstract: The main determinant factors with negative effect on development of the electronic libraries (EL) and use of electronic scientific and technical information in general in global network are considered from system positions. Approaches to creation of the mechanism of semantic navigation and search of knowledge in information networks are shown. The most priority and relevant directions and tasks in the field of increase in efficiency of development and use of EL are designated. It is noted that the organization of works on creation of EL is the cross-disciplinary task demanding consolidation of forces of experts in target subject domains in the field of information technologies, management and the right.

Keywords: electronic libraries; information environment; risks; information resources; network navigation and search; analytical post-processing; webometrics; information infrastructure

For citation: Syntyurenko O. V. Electronic libraries: problems and priority directions of development. In: *World-Class Scientific Publication – 2019: Strategy and Tactics of Management and Development: Proc. 8th Int. Sci. & Pract. Conf., Moscow, April 23–26, 2019*. Ekaterinburg: Ural University Press, 2019, pp. 101–107. DOI: 10.24069/konf-23-26-04-2019.15.

Становление цифровой экономики, развитие глобальной информационной инфраструктуры, цифровая трансформация экономического пространства активизируют осознание доминирующей роли информационных ресурсов (ИР) и технологий в процессах мирового экономического и социального развития. Осознавая важность

проблематики для развития страны, Правительство РФ разработало и в июле 2017 г. утвердило программу развития цифровой экономики до 2024 г. [1].

Электронные библиотеки (ЭБ), являющиеся относительно новым информационным ресурсом, междисциплинарным по своей сути, стали частью практически всех сфер деятельности. Технологии ЭБ становятся все более необходимыми для развития науки, образования, экономики, здравоохранения и т.д. Как область научных исследований проблематика ЭБ является составной частью процесса развития информационных технологий и охватывает широкий спектр вопросов теоретического и прикладного характера. Динамичность информационной среды, изменения социально-экономических и научно-технических факторов информатизации требуют системного подхода при изучении, анализе и классификации проблем развития и эффективности использования ЭБ. Перспективное развитие ЭБ в рамках развития национальной информационной инфраструктуры требует системных решений научно-методологических, правовых и других вопросов, не ограничиваясь только лишь технологическим подходом.

Термин «электронные библиотеки» используется специалистами в разных предметных областях и информационных системах. При этом под понятие «электронных библиотек», с некоторой долей условности, подпадают:

- 1) электронные издания и журналы (существуют как самостоятельно, так и в различных коллекциях);
- 2) различные электронные архивы; репозитории;
- 3) электронные коллекции визуализированных материальных объектов, депозитарии;
- 4) электронные библиотечные системы агрегаторов и производителей электронного контента;
- 5) электронные базы данных результатов выполненных исследований и разработок исследовательских организаций и научных фондов;
- 6) ресурсы для e-learning (электронные учебники и учебные пособия);
- 7) электронные каталоги различных видов ресурсов (в том числе по производимой и потребляемой промышленной продукции);
- 8) крупные национальные электронные библиотеки в глобальной сети.

Следует отметить, что электронная корпоративная информация ежегодно растет примерно на 40%. При этом, по экспертным оценкам, количество корпоративных пользователей Интернета возрастет до 2,7 млрд в 2019 г. по сравнению с 1,6 млрд в 2011 г.

С системных позиций рассмотрим основные факторы-детерминанты, отрицательно влияющие на развитие ЭБ, и использование электронной научно-технической информации в целом в глобальной сети.

1. Авторское (и патентное) право в его современной интерпретации является ведущим фактором противодействия развитию ЭБ и информационной свободе в Интернете [2]. Здесь следует отметить, что в 1982 г. специалисты Международной федерации библиотечных ассоциаций (IFLA) пришли к выводу, что действующие положения авторского права, вполне достаточные в рамках интеллектуального производства, не

в состоянии регулировать предоставление информации и информационных продуктов и услуг в виде баз данных. В 80х гг. прошлого века для программного обеспечения (операционная система UNIX) и некоторых других результатов интеллектуального труда получил развитие такой институт, как “публичный”, “бесплатный” и “условно бесплатный” объект (*public domain, freeware, shareware*), который распространяется бесплатно или за плату, которую готов заплатить потребитель при сохранении неимущественных авторских прав. Развитые страны в последнее десятилетие обращают все большее внимание на проблему открытости и доступности результатов научных исследований. Открытость в перспективе позволит повысить прозрачность науки, сократить нерациональные затраты, существенно снизить издержки от финансирования дублирующих исследований. По плану Европейской комиссии в рамках 7-летней программы Horizon 2020 (бюджет 80 млрд дол.) уже к 2016 г. 60% всех публикаций европейских ученых, проводящих свои исследования за государственный счет, должен размещаться в журналах открытого доступа. Некоторые страны (Австралия, Великобритания, США, и др.) уже сейчас на самом высоком уровне занимаются решением этой проблемы. Дания планировала реализовать открытый доступ в сети Интернет 80% научных статей к 2017 г. В России реализуется пилотный проект КиберЛенинка, где по лицензионному договору размещено около 500 научных журналов открытого доступа. В списке имеющихся в eLIBRARY журналов открытого доступа – 3108 наименований.

2. Факторы-детерминанты, обусловленные финансовыми ограничениями развития ЭБ и научных электронных информационных ресурсов в целом. В развитых странах существуют национальные программы создания электронных библиотек, финансируемые из госбюджета. Это, например, проект “Digital Library” по оцифровке исторических документов из фондов Библиотеки Конгресса США. В Германии данное направление охвачено проектом Global Info. В Японии стоимость проекта Электронной библиотеки 21-го века составляет 500 млн дол.

Объем средств, выделяемых на развитие системы НТИ, находится в прямой зависимости от общего объема финансирования исследований и разработок. В России общая сумма ассигнований федерального бюджета на научную деятельность в 2014 г. составила 437,3 млрд руб., или 0,61% к объему ВВП¹. Для сравнения – общий объем расходов Германии на НИОКР составляют примерно 2,85% ВВП (>70 млрд евро). При этом на развитие системы НТИ Германии выделяется около 1,5% от общего объема инвестиций в науку². Следует отметить, что в последние 5–10 лет в стране нарастает миграция к подписке на зарубежную научную периодику в электронной форме. Увеличиваются затраты на электронную подписку и их доля в общем бюджете на комплектование (особенно эта тенденция прослеживается в российских университетах).

3. Ограничения, обусловленные наличием языкового барьера. В процессе информатизации научно-технической деятельности и развития ЭБ все более возрастает значение лингвистических аспектов использования научных информационных ресурсов. Задача минимизации негативных факторов, обусловленных существованием

¹ По данным на 2015 г. ВВП России примерно соответствует ВВП Испании.

² Автор не смог найти по России аналогичные статистические данные.

языкового барьера, является частью проблемы формирования новой информационной культуры (по экспертным оценкам каким-либо иностранным языком в России владеет не более 15% инженеров и специалистов). Значимый фактор – все более широкое распространение в мире английского языка, который в настоящее время становится не только самым распространенным языком международного общения, но также базовым языком глобального информационного общества. На английском языке сегодня представлено около 90% научных знаний и более 80% электронных баз данных. В обозримом будущем, благодаря развитию процессов глобализации и информатизации общества, тенденция к дальнейшему распространению английского языка не только сохранится, но и будет нарастать. Определенным фактором негативного характера является также опубликование ценных документов, например описаний изобретений к патентам, на других, в том числе на редких, языках. Следует отметить, что задачи минимизации ограничений, обусловленных наличием языкового барьера, довольно успешно решаются путем все более широкого внедрения в практику поиска и обработки информации систем автоматического и автоматизированного перевода. Действующие системы машинного перевода ориентированы на конкретные пары языков (например, английский и русский). Качество машинного перевода зависит от объема словаря, объема информации, приписываемой лексическим единицам, от тщательности составления и проверки алгоритма анализа и синтеза, от эффективности программного обеспечения [3–5].

4. Темпоральные детерминанты (временные особенности) ЭБ и научных электронных журналов. Электронные научные журналы представляют собой вид сетевых информационных ресурсов, а учитывая высокую динамику сетевого информационного пространства, возникает высокая степень вероятности утраты или искажения данных публикаций. Достаточно много фактов прекращения существования электронного информационного ресурса (по тем или иным причинам). Очевидно, что без решения задачи создания и сохранения архивов научных публикаций электронные журналы не могут стать авторитетным источником научной информации. Следует отметить, что отчасти проблема сохранности во времени электронных научных журналов в России решается путем создания электронного архива (НТЦ «Информрегистр»), обеспечивающего надежное долговременное хранение юридически значимых электронных копий публикаций в электронных журналах. Такой электронный архив играет роль резервного информационного фонда, что, во-первых, гарантирует неизменность опубликованных материалов, во-вторых, обеспечивает их аккумуляцию и, в-третьих, позволяет сохранять полученные результаты вне зависимости от судьбы самого журнала как электронной системы (сбои технических средств, смена адресов, прекращение существования и т.п.).

5. Факторы-детерминанты развития ЭБ, обусловленные несовершенством механизма смысловой навигации и поиска знаний в информационных сетях. По оценкам специалистов, в информационно-поисковых системах (ИПС) по науке и технике коэффициент полноты поиска обычно составляет 50–60%, а коэффициент точности – 40–50%. В полнотекстовых поисковых системах сети Интернет эти показатели намного ниже. Увеличение этих параметров на каждые 10% требует практически удвоения

сложности лингвистических и программных средств. Причем если показатель точности дается пользователю, так сказать, в непосредственном восприятии, то потери релевантных документов остаются незамеченными, а вычисление показателя полноты требует выполнения специальной работы. Очевидно, что обеспечить приемлемую полноту без специальных лингвистических средств и методов практически невозможно. Именно поэтому большинство реально работающих документальных ИПС строится на основе вербальных информационно-поисковых языков с простой грамматикой.

В современной информатике отсутствуют методы универсального представления научного знания, которые позволяли бы интегрировать знания, накопленные в различных информационных ресурсах. Наиболее распространенным инструментом описания и упорядочения знаний являются тематические библиографические классификационные системы. Однако разные ресурсы используют разные, не сводимые друг к другу классификации. Потенциально это является фактором-детерминантом неэффективного использования научной информации, что приводит к невозможности интегрировать разрозненные источники знания для получения полной картины, которая необходима для продвижения научных знаний, в том числе в практику инновационных решений (и трансфера технологий). Следует отметить, что в настоящее время теория научно-технической информации не располагает методами индустриальной интеграции знаний, представленных в разнородных источниках. В ВИНТИ РАН ведутся исследования, в ходе которых применяются как интеллектуальные, так и автоматические методы анализа содержания классификационных систем и их соотношений. Интеллектуальный анализ состоит в определении теоретико-множественных отношений наполнения классификационных рубрик, включая систематическое определение между ними отношений тождества, включения, пересечения, либо отсутствия значимых смысловых связей. Новизна этой методики заключается именно в систематичности определения полного набора отношений между наполнением (денотатами) классификационных рубрик. Если в существующих классификациях систематически работают только отношения включения множеств, а отношения пересечения отмечаются только эпизодически (еще реже – отношения тождества), то в составе онтологии классификаций НТИ будет определено то или иное соотношение наполнений для всех рубрик рассматриваемых классификаций [6; 7]. Автоматический анализ соотношения тематики рубрик будет проводиться в соответствии с алгоритмом вывода прямых соответствий двух рубрик на основе имеющихся заданных отношений к третьей рубрике. Такой алгоритм разработан впервые. Новое технологическое решение позволит осуществлять:

- переход от одного ресурса к другому по определенной тематике;
- объединение знаний по определенной тематике, содержащихся в различных информационных ресурсах;
- сравнение тематического наполнения различных информационных ресурсов;
- определение областей компетенции агентов информационной деятельности (организаций, исследователей, менеджеров).

Исследования в ВИНТИ выполняются при поддержке Российского фонда фундаментальных исследований (грант № 17-07-00153).

С учетом вышерассмотренных проблем обозначим наиболее приоритетные и актуальные направления и задачи в области повышения эффективности разработки и использования ЭБ:

1. Значительное (более чем в два раза) увеличение темпов оцифровки информационных ресурсов, особенно в научно-технической сфере.
2. Разработка и широкое внедрение эффективных методов смысловой навигации и поиска знаний в информационных сетях.
3. Минимизация временных факторов-детерминат использования ЭБ и научных электронных журналов.
4. Минимизация факторов-детерминант развития ЭБ, обусловленных наличием языкового барьера.
5. Расширение применения вебметрических систем и исследований в пространстве ЭБ. Вебметрические системы и исследования, основывающиеся на статистическом анализе основных показателей цифровых ИР, являются относительно новым научным направлением и эффективным инструментом совершенствования ЭБ [8]. Аналитическая постобработка вебметрической информации обеспечивает выявление слабых и сильных сторон ЭБ, веб-сайтов научных организаций, анализ пользовательской аудитории, формирование сопоставительных рейтинговых оценок.
6. Мультипликативная аналитическая постобработка научно-технической и технико-экономической информации ЭБ и баз данных с использованием методов наукометрии, эконометрии и многомерного анализа данных.
7. Минимизация различного рода рисков разработки и функционирования электронных библиотек [9].

Заключение

1. Быстрое развитие электронных библиотек обуславливается несколькими факторами, среди которых: стремительное развитие информационно-телекоммуникационных технологий, лавинообразный рост информационных ресурсов в глобальной сети, увеличение количества документов, существующих только в электронной форме, рост темпов информатизации научно-технической сферы деятельности. Электронные библиотеки становятся одним из доминирующих факторов, формирующих современную информационную среду 21-го века. Создание крупных и доступных национальных электронных библиотек способствует более эффективному использованию информации, развитию науки и техники, экономики в целом.

2. Определяющее значение в происходящей трансформации экономики приобретают исследования, разработки, трансфер технологий, что требует создания современной информационной инфраструктуры и эффективной системы комплексной информационной поддержки цикла «исследование – разработка – производство», а также всей системы специального образования. Основным и важнейшим фактором является развитие электронных сетевых информационных ресурсов.

3. В комплекс приоритетных мероприятий по развитию информационной инфраструктуры научно-технической и промышленной сферы должны входить организационно-финансовая и методологическая поддержка контентного и технологического

развития ЭБ и работ по оцифровке информационных фондов научно-технических библиотек.

4. Необходимо особо отметить, что организация работ по созданию ЭБ подразумевает и требует широкой консолидации сил специалистов в целевых предметных областях, в области информационных технологий, управления, права и образования. Это – междисциплинарная задача. Решение этих проблем требует также значительных финансовых и интеллектуальных (кадровых) инвестиций. К созданию ЭБ должны привлекаться специализированные организации и творческие коллективы, способные профессионально и качественно выполнить все необходимые этапы разработки.

Список литературы

1. Национальная Программа «Цифровая экономика Российской Федерации»: утверждена распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р. Режим доступа: <http://static.government.ru/media/files/9gFM4FHj4PsB79I5v7yLVuPgu4bvR7M0.pdf>

2. Сюнтюренко О. В. Факторы-детерминанты неэффективного использования информационных ресурсов в научно-технической деятельности. *Научно-техническая информация. Сер. 1.* 2017;(7):1–12.

3. Brynjolfsson E., McAfee A. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York: Norton & Company; 2016. 320 p.

4. Дроздова К. А. Машинный перевод: история, классификация, методы. В: *Филологические науки в России и за рубежом: материалы III Междунар. науч. конф., г. Санкт-Петербург, июль 2015 г.* СПб.: Свое издательство; 2015. С. 139-141. Режим доступа: <https://moluch.ru/conf/phil/archive/138/8497/> [Дата обращения: 28.12.2018].

5. Колганов Д. С., Данилов Е. А. Обзор аналитической, статистической и нейронной технологии машинного перевода. *Международный студенческий научный вестник.* 2018;(3-2). Режим доступа: <http://eduherald.ru/ru/article/view?id=18262> [Дата обращения: 28.12.2018].

6. Antopol'skii B., Beloozerov V. N., Kalenov N. E., Shaburova N. N., Yakshin M. M. The Development of a Semantic Network of Keywords Based on Definitive Relationships. *Scientific and Technical Information Processing.* 2017;44(4):261–265. DOI: [10.3103/S0147688217040062](https://doi.org/10.3103/S0147688217040062)

7. Белоозеров В. Н., Антошкова О. А., Шапкин А. В., Щуко Ю. Н. Разработка семантической сети библиографических классификационных систем. В: *Наукометрия, библиометрия, открытые данные и публикации в науке: материалы девятнадцатой международной конференции и выставка «LIBCOM-2018», г. Суздаль, 26–30 ноября 2018 г.* Режим доступа: <http://www.gpntb.ru/libcom2018/doc/010.pdf> [Дата обращения: 28.12.2018].

8. Булычева О. С., Сюнтюренко О. В. Концептуальные положения и предпосылки создания вебметрической системы цифрового пространства библиотек. В: *Сборник Президентской библиотеки им. Б. Н. Ельцина. Сер. «Электронная библиотека».* Вып. 8. СПб.: Президентская библиотека; 2018. С. 19–31.

9. Булычева О. С., Сюнтюренко О. В. Электронные библиотеки: перспективы и риски развития библиотек. В: *Сборник научных трудов Президентской библиотеки им. Б. Н. Ельцина. Сер. «Электронные библиотеки».* Вып. 6: Интегрированные цифровые ресурсы: организационно-технологические и научно-методические основы развития. СПб.: Президентская библиотека; 2015. 271 с.

Информация об авторах

Сюнтюренко Олег Васильевич – доктор технических наук, профессор, ведущий научный сотрудник Научно-методологического отделения, Всероссийский институт научной и технической информации Российской академии наук (ВИНИТИ РАН), г. Москва, Россия; e-mail: olegasu@mail.ru.

Information about the author

Oleg V. Syuntyurenko – Doctor of Sciences (Engineering), Professor, Leading Researcher of Scientific and Methodological Department, All-Russian Institute for Scientific and Technical Information of the Russian Academy of Sciences (VINITI RAS), Moscow, Russia; e-mail: olegasu@mail.ru.